A Computer Program for Objective Point Symmetry Classifications of Electron Diffraction Spot Patterns with Apparent Hexagonal or Rectangular-Centered Lattice Metric

Lukas von Koch[1*,2] and Peter Moeck[2*]
[1] Westside Christian High School, Portland, Oregon, USA, *lukasvonkoch@gmail.com
[2] Department of Physics, Portland State University, Portland, Oregon, USA, *pmoeck@pdx.edu

A Python program has been written to perform complementary calculations to the recently developed information-theoretic approach to crystallographic symmetry classifications/ quantifications [1-4] in two dimensions (2D) from digital atomic or molecular resolution transmission electron and scanning probe microscope images. This program is briefly described below. An example of its usage is given in the form of selected results from the classification/quantification of the experimental electron diffraction pattern in Fig. 1A. The information-theoretic symmetry quantifiers are briefly compared with the traditional (crystallo-graphic) R_{sym} values [5]. For further results of analyses of the diffraction pattern in Fig. 1A, see our free on-line material [6,7].

For the purpose of the classification/quantification, digital input data consist in [1-4,6,7] of the pixel-wise sums of approximately Gaussian distributed noise and a strictly 2D periodic unknown underlying signal. Structural defects within the crystal or on the crystal surfaces, instrumental image recording noise, slight deviations from zero-crystal-tilt conditions in transmission electron microscopy, and small inaccuracies in the algorithmic processing of the digital data all contribute to a single generalized noise term. The plane symmetry group and projected Laue class [1,3] (or 2D Bravais lattice type [4] and 2D point symmetry group [6,7]) that are "anchored" to the least broken symmetries [1] are identified as genuine in the presence of generalized noise. More severely broken symmetries that are not anchored in this sense are identified as pseudo-symmetries.

The electron crystallography program CRISP/ELD 2.1 [5] was used in its default setting for the extraction of the electron diffraction spot intensities. The lattice parameters $a = 12.46 \pm 0.2$ Å, $b = 12.41 \pm 0.2$ Å, $\gamma = 119.5 \pm 1.0°$ were obtained with this program from the pattern in Fig. 1A. (This is an oblique lattice metric that is within error bars hexagonal.) Program results for a hexagonal indexing of this pattern are given as a screenshot in Fig. 1B.

The extracted diffraction spot intensities with primitive (hexagonal) Laue indices were exported as *.hke files and read into the program for analysis. Plain text files such as *.hke, *.cif, etc. typically contain information that is unnecessary for our analysis (i.e. more than the three columns mentioned in the following sentence or lines of text at the top of the file). Both of these issues are avoided by the program as it allows a user to tell it which lines of the file they wish to delete as well as which columns of the file represent h, k, and observed spot intensities. These three values must consistently occupy the same columns in each line of the input text file. Once settings are inputted by the user and a suitable data file has been read in, the program adds missing spot intensities and completes a number of calculations instantaneously. The program output consists of two result tables. Figures 1B and 2 are screenshots of the first of these tables. (A table with the confidence levels [1,2] of higher symmetric symmetry classifications is the second output). The data columns of the first result table are as follows: sums of normalized squared residuals, geometric Akaike Information Criteria (G-AIC) values, likelihoods to be the Kullback-Leibler [8] (K-L) best geometric model of the experimental data, geometric Akaike

weights [1-3], and traditional R_{sym} values. The third, fourth, and fifth columns of these tables are of particular importance to our study (and our scientific progress with respect to reliance on R_{sym} values for classifications of electron diffraction spot patterns into point symmetry groups).
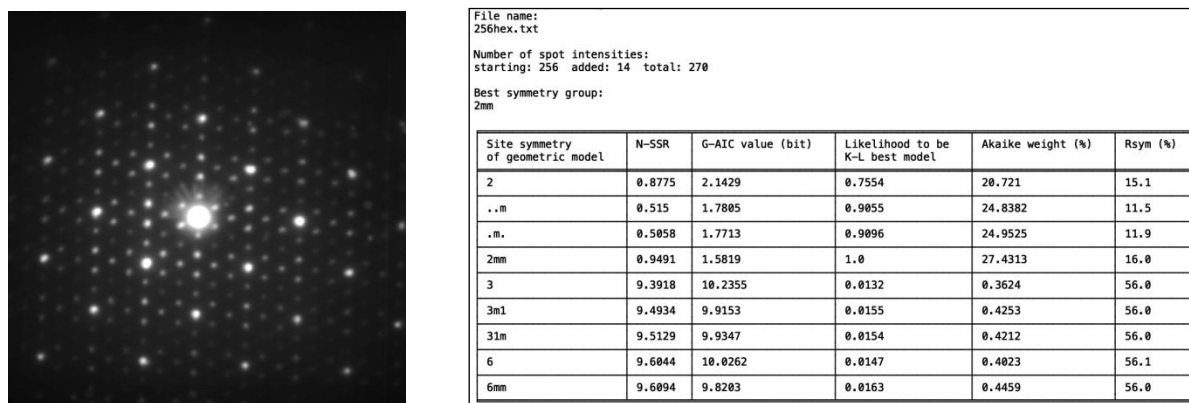
File name:
256hex.txt

Number of spot intensities:
starting: 256 added: 14 total: 270

Best symmetry group:
2mm

Site symmetry of geometric model	N–SSR	G–AIC value (bit)	Likelihood to be K–L best model	Akaike weight (%)	Rsym (%)
2	0.8775	2.1429	0.7554	20.721	15.1
..m	0.515	1.7805	0.9055	24.8382	11.5
.m.	0.5058	1.7713	0.9096	24.9525	11.9
2mm	0.9491	1.5819	1.0	27.4313	16.0
3	9.3918	10.2355	0.0132	0.3624	56.0
3m1	9.4934	9.9153	0.0155	0.4253	56.0
31m	9.5129	9.9347	0.0154	0.4212	56.0
6	9.6044	10.0262	0.0147	0.4023	56.1
6mm	9.6094	9.8203	0.0163	0.4459	56.0

Fig. 1 (**A-Left**) Central part of an experimental electron diffraction spot pattern from [5]. The quasi-horizontal mirror line is *..m* and its quasi-vertical counterpart is *.m.* for a rectangular-centered indexing. (**B-Right**) Screenshot of the first result table of the analysis program for a hexagonal (primitive) indexing.

The geometric model with point symmetry group *2mm* is the K-L best model of the experimental data in Fig. 1A as it has the highest geometric Akaike weight. The average confidence level for preferring point symmetry *2mm* over its three maximal subgroups as information-theoretic point symmetry classification of the pattern in Fig. 1A is 38.83%. One can deduce from the table in Fig. 1B that the electron diffraction pattern in Fig. 1A does not feature a hexagonal point symmetry group [7]. Thus, the pattern should be re-indexed for the rectangular-centered Bravais lattice type and the analysis repeated, leading to the screenshot in Fig. 2. This repeat confirms the ability of the program as its numerical results cannot depend on the labels of the electron diffraction spots. (CRISP/ELD extracted intensities are naturally expected to be slightly different for the two different indexings of the diffraction pattern in Fig. 1A). The geometric Akaike weights for point symmetry group *2mm* in Figures 1B and 2 differ by only 0.04%. This weight is, thus, a more robust and sophisticated quantifier than the average confidence level for ascent to *2mm* from its three maximal subgroups. The latter came to 40.08% for the rectangular-centered indexing of the diffraction pattern in Fig. 1A.

```
File name:
256rec.txt

Number of spot intensities:
starting: 256  added: 2  total: 258

Best symmetry group:
2mm
```

Site symmetry of geometric model	N–SSR	G–AIC value (bit)	Likelihood to be K–L best model	Akaike weight (%)	Rsym (%)
2	0.8736	2.1606	0.7589	20.3208	15.4
..m	0.5229	1.81	0.9043	24.2149	11.8
.m.	0.534	1.8211	0.8993	24.0807	12.1
2mm	0.9653	1.6088	1.0	26.7773	16.3
3	7.7665	8.6245	0.03	0.8023	52.0
3m1	7.8606	8.2896	0.0354	0.9485	52.1
31m	7.8675	8.2965	0.0353	0.9453	52.1
6	7.9552	8.3842	0.0338	0.9047	52.2
6mm	7.9584	8.1729	0.0376	1.0055	52.1

Fig. 2. Screenshot of the first result table of the program for rectangular-centered indexing.

Electron diffraction spot pattern based electron crystallography is set to benefit from the new classification/quantification method and computer program for objective crystallographic symmetry classifications and quantifications. Current symmetry classification practices in that field are made on the basis of relatively low R_{sym} values [5] that can only be interpreted subjectively. A consequence is that crystal structures are occasionally solved and refined in the wrong space groups, see appendices of [3].

The space group of the crystal under investigation might be overestimated or underestimated. If it is underestimated, one does not use the available experimental data optimally and ends up publishing a result that could have been more accurate by averaging over smaller and more numerous asymmetric units in direct and/or reciprocal space. When the crystal's space group has been overestimated, e.g. twinning by pseudo-merohedry has been overlooked, one averages over parts of the structure that are not genuinely related to each other by symmetry. In other words, the estimated asymmetric unit is too small and the derived crystal structure is incorrect.

Underestimation is, therefore, less harmful than overestimation as far as space group symmetry classifications are concerned. The geometric form of information theory [1-4] that was adapted here for the purpose of point symmetry classifications/quantifications of electron diffraction spot pattern is known to "err on the side of caution" in this respect. Over time and with better experimental data, the entries in the major crystallographic databases will become more accurate when objective crystallographic symmetry classifications [1-4] have been generalized to three dimensions and become mainstream [9].

References:

[1] P Moeck, Acta Cryst. A, vol. 78, 2022, in print, 23 pages, doi: 10.1107/S2053273322000845, (expanded version, 35 pages, arXiv: 2108.00829, Jan. 4, 2022).
[2] A Dempsey and P Moeck, 74 pages, arXiv: 2009.08539, Dec. 15, 2020.
[3] P Moeck, Symmetry **10** (2018), p. 133.
[4] P Moeck in "Microscopy and Imaging Science: Practical Approaches to Applied Research and Education", ed. A. Méndez-Villas, (Badajoz: FORMATEX, 2017) p. 503; arXiv: 2011.13102v2.
[5] X Zou, S Hovmöller, and P Oleynikov, Electron Crystallography: Electron Microscopy and Electron Diffraction, (Oxford University Press, 2011).

[6] P Moeck and L von Koch, arXiv: 2201.04789, 4 pages, Jan.-Feb. 2022.

[7] P Moeck and L von Koch, arXiv: 2202.00220, 4 pages, Feb. 2022.

[8] D. R. Anderson, Model Based Inference in the Life Sciences: A Primer on Evidence, New York, NY: Springer Science+Business Media, 2008.